

参赛密码 _____
(由组委会填写)

第十二届“中关村青联杯”全国研究生 数学建模竞赛

学 校 解放军空军预警学院

参赛队号 90047002

 1. 卢雨祥

队员姓名 2. 李 帅

 3. 李佳炜

参赛密码 _____
(由组委会填写)



第十二届“中关村青联杯”全国研究生 数学建模竞赛

题 目 移动通信中的无线信道“指纹”特征建模

摘 要：

论文针对无线多径信道的特征提取，以及无线信道指纹建模做了以下一些研究：

针对第一问，论文首先基于 MPOC 算法，求解了多径信道的单位脉冲响应，并且基于多径信道的单位脉冲响应，提取了多径数目、多径时延、最大多径展宽、主多径平均功率、主多径功率方差等待选特征；针对多普勒展宽，论文利用 FFT 分析了一组测试数据中都有信道样本同一时刻的数据，得到了多普勒频谱，并提取了多普勒展宽。在提取多径数目与多普勒展宽时，论文利用数据中较小的点求平均构造噪声基底作为检测门限，将信号衰减 15Db 后，过噪声门限作为有效数据，以此来防止噪声数据被作为有效数据使用。但是利用 MPOC 算法求解了多径信道的单位脉冲响应，并且基于上述峰值检测方法检测多径信道上的路径，一般都可以检测出多个路径，但除少数几个主峰以外，其余多径的增益系数很小，很难区分是多径还是主峰引起的旁瓣或噪声。所以基于这种方法

检测出的多径并不稳定。基于以上分析，论文提出在多径个数检测时，选取最大主峰值幅值 0.5 倍以上的峰值作为主多径。

针对信道指纹的评价指标，论文提出：对于三个场景各五个测试样本的“指纹”特征，利用 K-means 聚类进行分类，如果出现两个不同场景的测试样本特征聚为一类，那么认为特征指纹模型不合理；反之，认为特征选取合理。经过聚类分析，论文得出以下信道“指纹”特征：

主多径数	表征发射信号衰减较小的路径数
主多径平均功率	度量主多径对信号的平均衰减能力
多普勒扩展	反映多径信道时变性质的强弱

针对第二问，论文基于第一问提出的三个信道指纹特征，对问题二中的两组测试数据进行特征提取，将原有的 15 个特征指纹与新加入的测试指纹进行聚类，通过分类结果，判断第二问中的测试属于哪一个场景。得到结果如下：

场景一	Test1ForScene.mat
场景二	Test2ForScene.mat

针对第三问，论文采用由粗划分到精细划分的划分原则，首先对“Sample.mat”中的数据提取主多径数目，依据提取到的主多径，初步将“Sample.mat”划分为 3 段；在粗划分的基础上，以 100 个样本为步进提取多普勒展宽以及主多径功率特征，利用 K-means 算法对粗划分的三类分别进行细化分，最终得到了划分为 12 段。

针对第四问，类似于第二问的处理方法，分别提取两个测试样本的信道指纹，与第三问的各个场景特征做相关系数，找到相关系数最大的场景，认定该测试样本采样于该场景，最终发现测试信道一处在路段的 2.5m 到 21.7m，测试信道二处在路段的 47.5m 到 75m。

关键词：多径信道指纹、MPOC、FFT、K-means、主多径

一、问题引入

1.1 无线指纹

移动通信产业一直以惊人的速度迅猛发展，已成为带动全球经济发展的主要高科技产业之一，并对人类生活及社会发展产生了巨大的影响。在移动通信中，发送端和接收端之间通过电磁波来传输信号，我们可以想象两者之间有一些看不见的电磁通路，并把这些电磁通路称为无线信道。无线信道与周围的环境密切相关，不同环境下的无线信道具有一些差异化的特征。如何发现并提取这些特征并将其应用于优化无线网络，是当前研究的一个热点。类比人类指纹，我们将上述无线信道的差异化的特征称为无线信道“指纹”。无线信道“指纹”特征建模，就是在先验模型和测试数据的基础上，提取不同场景或不同区域内无线信道的差异化的特征，进而分析归纳出“指纹”的“数学模型”，并给出清晰准确的“数学描述”。

1.2 多径传播与时延扩展

文献[1]中提到：无线信道与有线信道最大的区别在于有线信道中信号是沿着特定的媒质通过单一路径到达接收端的，而无线信道在传输过程中，电磁波是在空间中自由传播的，接收机收到的信号是由很多经不同传播路径到达的信号构成的。这些不同路径的信号有着不同的传输相位，叠加在一起有的会互相增强，而有的会互相削弱，因此接收机接收到的信号就会产生衰落。由于传输时延不同，发射端发送的单个脉冲信号，在接收端可能会陆续收到多个脉冲信号或者是长度展宽的脉冲信号，这称为时延扩展。

1.3 多径效应与多普勒扩展

文献[2]中提到：由于移动台的运动会使得移动台和基站之间的无线信道呈现出时变性，时变性是指信道的传递函数是随时间而变化的，即不同的时刻发送相同的信号，在接收端收到的信号是不相同的。时变性在移动通信系统中的具体体现之一是多普勒频移(Doppler Shift)，即单一频率信号经过时变衰落信道之后会呈现为具有一定带宽和频率包络的信号，称为信道的频率弥散性(Frequency Dispersion)。

多普勒效应所引起的附加频率偏移称为多普勒频移，可以用下式表示：

$$f_d = \frac{v}{\lambda} \cos(\theta) = \frac{vf_c}{c} \cos(\theta) = f_{d\max} \cos(\theta) \quad (1-1)$$

多普勒频移与载波频率和移动台运动速度成正比。当移动台向入射波方向移动时，多普勒频移为正，即移动台接收到的信号频率会增加；如果背向入射波方向运动，则多普勒频移为负，即移动台接收到的信号频率会减小。

从时域来看，与多普勒频移相关的是相干时间，是信道传输函数的时间相关函数大于某个门限值的时间间隔统计平均值。若门限值设为 0.9，相干时间近似为：

$$(\Delta T_c) \approx \frac{1}{f_{d\max}}, \text{ 如果门限值设为 } 0.5, \text{ 相干时间近似为 } (\Delta T_c) \approx \frac{9}{16\pi f_{d\max}}。$$

相干时间就是指一段时间间隔，在此间隔内，两个到达信号有很强的幅度相关性。如果基带信号带宽的倒数，一般指符号宽度大于无线信道的相干时间，那么信号的波形可能发生变化，造成信号的畸变，产生时间选择性衰落，即快衰落；反之，如果符号的宽度小于相干时间，信号经历非时间选择性衰落，即慢衰落。

1.4 多径传播的数学模型

在无线信道中传输的信号都是带通的，因此发射的带通信号的复包络可以表示为：

$$s'(t) = \text{Re}[s(t)e^{j2\pi f_c t}] \quad (1-2)$$

(3-1) 式中是 $s(t)$ 是发射信号， f_c 是信号载频。

信号经过无线信道，会产生多径效应。所谓多径效应，即信号沿不同路径到达接收端。因为信号在多条路径上传播，信号接收信号可以看做发射信号以不同时间延迟，不同的衰减系数的线性组合。假设此时移动台与发射台处于相对静止，此时回波信号表示为：

$$r(t) = \sum_{i=1}^N a_i s'(t - \frac{x_i}{c}) \quad (3-2)$$

将 (3-1) 式代入 (3-2) 式得：

$$r'(t) = \sum_{i=1}^N a_i \operatorname{Re}\{s(t - \frac{x_i}{c}) \exp[j2\pi f_c(t - \frac{x_i}{c})]\} \quad (3-3)$$

式中 x_i 和 a_i 分别代表第 i 条路径的路径长度和反射系数, c 是光速, N 为多径数。

将 $r'(t)$ 写成复信号的形式:

$$r'(t) = \operatorname{Re}[r(t)e^{j2\pi f_c t}] \quad (3-4)$$

当移动台以速度 v 移动、接收信号入射方向与移动台运动方向的夹角为 θ_i , 则产生的多普勒频移 f_d 为:

$$f_d = \frac{v}{\lambda} \cos(\theta_i) \quad (3-5)$$

移动台运动产生的路径长度的变化量为:

$$\Delta x_i = -vt \cos(\theta_i) \quad (3-6)$$

因此每条径的频率都会产生变化, 在这种情况下接收信号可表示为:

$$\begin{aligned} r(t) &= \sum_{i=1}^N a_i e^{-j2\pi(\frac{x_i + \Delta x_i}{\lambda})} s(t - \frac{x_i + \Delta x_i}{c}) \\ &= \sum_{i=1}^N a_i e^{-j2\pi\frac{x_i}{\lambda}} e^{-j2\pi\frac{v}{\lambda}t \cos\theta_i} s(t - \frac{x_i}{c} + \frac{vt \cos\theta_i}{c}) \end{aligned} \quad (3-7)$$

与 $\frac{x_i}{c}$ 相比 $\frac{vt}{\cos\theta_i}$ 很小, 可以忽略, 并且令 $a_i' = a_i e^{-j2\pi\frac{x_i}{\lambda}}$, 则有:

$$\begin{aligned} r(t) &= \sum_{i=1}^N a_i' e^{-j2\pi\frac{v}{\lambda}t \cos\theta_i} s(t - \tau_i) \\ &= \sum_{i=1}^N a_i' e^{-j2\pi f_d t} s(t - \tau_i) \end{aligned} \quad (3-8)$$

$\tau_i = \frac{x_i}{c}$ 和 $f_{di} = \frac{v}{\lambda} \cos\theta_i$ 分别定义为第 i 条路径的时延和多普勒频率。由此可以看出, 多径每条信道都在发射信号上调制了复增益以及时延和多普勒频率。

二、模型假设以及基本符号说明

(1) 假设噪声是白噪声;

(2) 假设在信道样本内信道具有线性时不变的单位序列响应 $h(k)$, 而在信道样本间信道的单位脉冲响应具有时变性, 看做线性时变系统;

(3) 假设发射信号为一冲激信号, 且周期发射, 问题一、二中发射周期为 $\frac{2}{3}$ ms, 问题三、四中发射周期为 10ms。

三、多径特征提取以及评价指标

3.1 问题分析

3.1.1 多径特征分析

无线通信相较于传统有线通信，电磁波传输的路径并不是单一路径，而是具有多径效应。多径效应顾名思义，就是信号传输路径有多条，故信号到达接收端的时刻不同，将引起时延扩展。而由于信号在空间中散射（直射、反射、绕射）形式以及路径长度的不同，各条路径对信号的衰落也不同。而由于移动台的运动会使得移动台和基站之间的无线信道呈现出时变性，又会引起接收信号的多普勒扩展。所以，论文认为，可以提取多径数目、多径时延展宽、多普勒展宽信息、多径增益系数等信息，建立备选无线信道特征。而要提取信道特征，就必须求解出多径信道的冲激响应 $h(k,n)$ 。

3.1.2 多径信道的冲激响应

问题中，发射端发射了一个冲激信号，经过发射接收端滤波器与多径信道的调制，得到接收信号。即

$$r[k,n] = \sum_{i=0}^{M-1} h[k-m,n]g[m] + u[k,n] \quad (3-1)$$

其中，“ n ”表示测试的样本标识，对应测试时刻，这里假设共有“ N ”个样本；“ $h_l[n]$ ”表示“ n ”时刻第 l 条路径上的信道系数，通常是复数；“ $\tau_l[n]$ ”表示“ n ”时刻第 l 条路径延迟的样点数；“ $u[k,n]$ ”表示“ n ”时刻第“ k ”个测试样本上引入的复高斯白噪声；“ $r[k,n]$ ”表示“ n ”时刻单位脉冲依次经发送滤波器、信道和接收滤波器后的实际接收信号，是 k,n 的二元函数。显然，我们可以从“ $r[k,n]$ ”中获取完整的时变信道“ $h[k,n]$ ”。

根据问题假设（2），在信道样本内信道具有线性时不变的单位序列响应 $h(k)$ ，此时接收信号 $r[k,n]$ 可以看做多径信道冲激响应 $h[k,n]$ 与 $g[k]$ 的卷积，即：

$$r[k] = \sum_{m=0}^{M-1} h[k-m]g[m] + u[k], k = 0,1,...,K-1 \quad (3-2)$$

其中，“ M ”为滤波器的长度，即“ $g[k]$ ”的样点数。根据卷积的交换律，接收信号 $r[k,n]$ 可以看做输入信号 $g[k]$ 通过冲激响应为 $h[k,n]$ 的系统产生的，也可

以看做输入信号 $h[k,n]$ 通过冲激响应为 $g[k]$ 的系统产生的。

3.2 多径冲击响应的求取

根据 3.1.2 中的分析，这里将 $r[k,n]$ 看做输入信号 $g[k]$ 经过冲激响应为 $h[k,n]$ 的系统产生的。而已知输入信号 $g[k]$ 、系统输出 $r[k,n]$ 估计系统冲激响应经典的最小二乘估计、反卷积以及 MPOC 算法、MUSIC 算法、EM 算法、WRELAX 算法等高分辨算法。由于最小二乘估计、扩频滑动相关法等传统方法实验估计分辨力较低，论文基于分辨力较高的 MPOC 算法求解了多径信道的冲激响应 $h[k,n]$ 。考虑到系统中有加性白噪声的干扰，论文采用扩频滑动相关法（等效于匹配滤波器）进行了降噪处理。

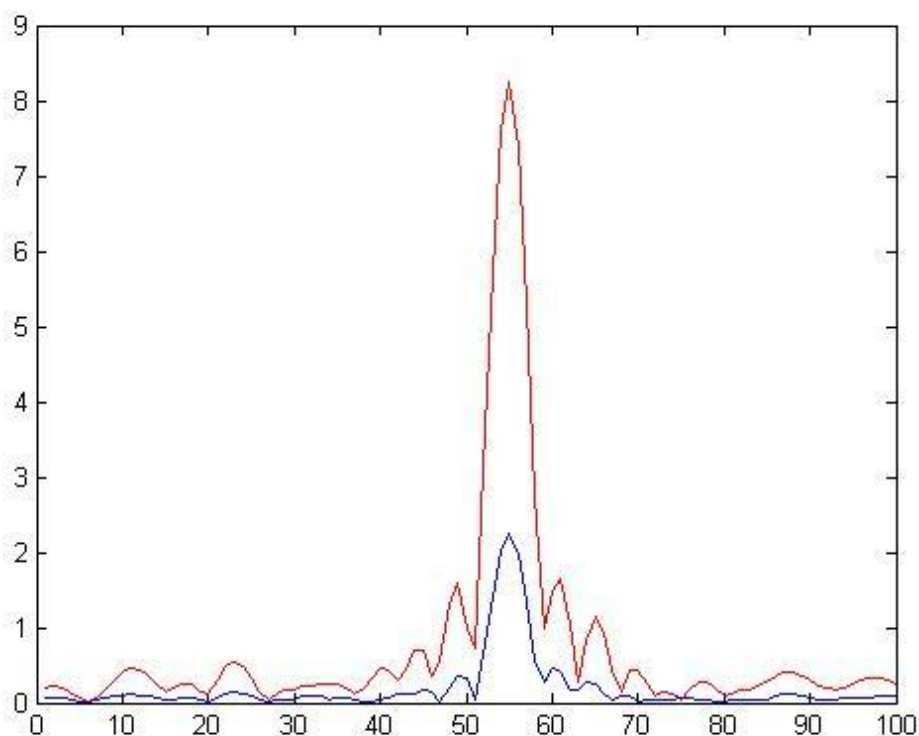


图 3-1

如图 3-1，接收信号经过频滑动相关处理后，信号主峰能量得到明显加强，说明达到了抑制噪声的效果。

3.2.1 基于扩频滑动相关法降噪处理

系统输入信号为 $g[k]$ ，输出信号为 $r[k,n]$ ，分别对 $g[k]$ 做傅里叶变换（Fourier Transform）得到 $g[k]$ 、 $r[k,n]$ 的频谱 $G[f]$ 、 $R[f,n]$ 。扩频滑动相关

法处理就是将 $g[k]$ 反转作为滤波器的系数与输出信号 $r[k,n]$ 进行卷积，这种处理等效于求 $g[k]$ 与 $r[k,n]$ 的相关函数，即：

$$rf[k] = g[-k] * r[k] \quad (3-3)$$

$rf[k]$ 为扩频滑动相关法后输出。

3.2.2 基于 MPOC 算法求解多径信道冲激响应

南洋理工的 Yuan Zhou, Yong Liang Guan 等人提出的 MPOC (Modified Phase-Only Correlator) 算法是一种直接利用频域数据估计信道的冲激响应，从而对接收信号进行多径时延估计，与传统反卷积较为类似，但通过降噪处理，提升了多径估计的准确性，相对于传统扩频滑动相关法，具有更高的时延分辨力。

信号源产生的信号 $s(t)$ 后，经过多径信道后到达接收端，信号在每条路径经历了不同的时延和衰减，还会受到噪声的影响，因而接收信号可以表示为：

$$r(t) = \sum_{i=1}^M h(i)g(t - \tau_i) + n(t) \quad (3-4)$$

$n(t)$ 表示噪声， M 表示信道多径数， a_i 和 τ_i 分别表示第 i 条径的衰减因子和时延，将式(3-4)表示成频域形式：

$$R(f) = \sum_{i=1}^M h(i)G(f)e^{-j2\pi f\tau_i} + N(f) \quad (3-5)$$

两边同时除以 $G(f)$ ：

$$H(f) = \frac{R(f)}{G(f)} = \sum_{i=1}^M h(i)e^{-j2\pi f\tau_i} + \frac{N(f)}{G(f)} \quad (3-6)$$

式(3-3)的第一项就是多径信道的频率响应，理论上讲对第一项做逆傅里叶变换就可以得到信道的脉冲响应，从而得出各条路径的延迟。第二项是噪声分量。由于 $G(f)$ 是一个带宽为 $246.2 M$ 的低通信号，在带外很小， $\frac{N(f)}{G(f)}$ 的值会变得很大，会放大噪声的影响。这也成为影响这种算法性能的主要因素。

去掉相除结果中的异常值，再进行 IFFT 运算，就得到了信道的时域脉冲响应。论文中，将反卷积得到的 $H(f)$ 出现在 $R(f)$ 通带外的奇异值都进行置 0 处理，再利用 $H(f)$ 进行傅里叶逆变换得到多径信道的冲激响应 $h(k,n)$ 。

3.3 特征提取

3.3.1 基于 MPOC 算法提取多径参数

不同多径的延时不同，在一个信道样本得到的冲击响应中，如果在时延上有多个峰值，即可认为该样本有多个多径。因此，判断多径数目首先要对每个样本进行峰值检测。在一个信道上大量的样本数据，由于检测误差的存在、分辨精度限制以及各参数具有时变性，这些样本的峰值数目并不完全一致，需要对峰值信息进行综合处理，从而确定该信道的多径数目。

下面先阐述样本上峰值检测方法。

信道的多径数目与时延特性都需要对功率延迟分布进行分析获得。考虑到信道中噪声的影响，为了保证后续统计参数的准确性和提升数据池的精度，数据处理时需要选取一个判决阈值来区分噪声和系统中的有效多径，进而对数据样本进行筛选。

基本的峰值衰减法确定判决阈值，进行样本筛选时，首先需要确定的两个基本参数是：功率延迟分布的峰值和样本的动态噪底。功率延迟分布的峰值可以通过峰值搜索得到。样本的动态噪底可以通过对信道噪声大小的近似估计得到，计算时，首先确定功率延迟分布中的低幅值点集合（所有采样幅值中最小的 20%）；其次剔除集合中 25% 的最高点和 25% 的最低点（类似中值滤波），并更新集合数据；最后对新集合中的样本点的幅值进行平均，作为动态噪底。峰值衰减算法要求样本功率延迟分布的峰值大于动态噪底 18dB，并将功率延迟分布峰值衰减 18dB 后的门限值作为判决阈值，这样有效样本的判决阈值和动态噪底之间的动态范围应该不小于 3 dB。最后，剔除不符合要求的样本，多径检测公式为：

$$20 \times \log_{10} h(k) - 15 > gate$$

$gate$ 为噪声均值门限， $h(k)$ 为功率延迟分布。

检测到多径后，利用每一个多径对应的下标，计算多径时延： $\tau_i = 65 * k$ ， τ_i 为多径时延， k 为检测到多径的位置（点数）。

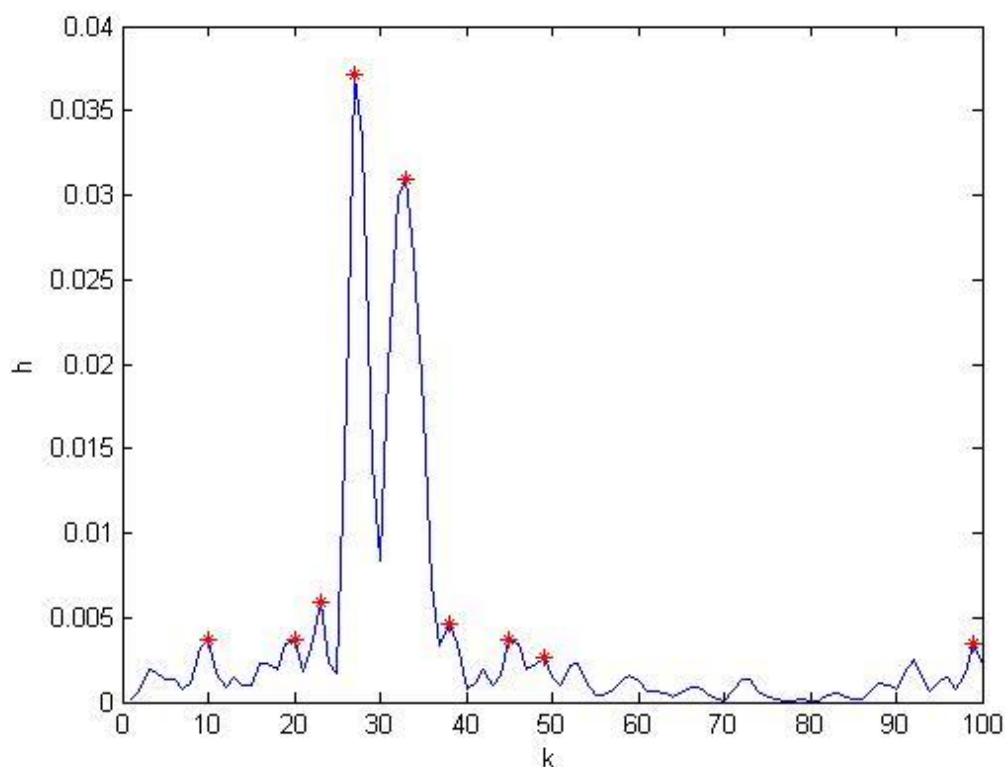


图 3-2 基于 MPOC 算法多径信道的单位脉冲响应的求解

如图，利用 MPOC 算法求解了多径信道的单位脉冲响应，并且基于上述峰值检测方法检测了多径信道上的路径。但是，从图发现多径信道一共检测出 9 个，但除了两个主峰以外，其余多径的增益系数很小，很难区分是多径还是主峰引起的旁瓣或噪声。所以基于这种方法检测出的多径并不稳定。基于以上分析，论文提出在多径个数检测时，选取最大峰值幅值 $\frac{1}{2}$ 以上作为主多径，如图 3-2 所示：

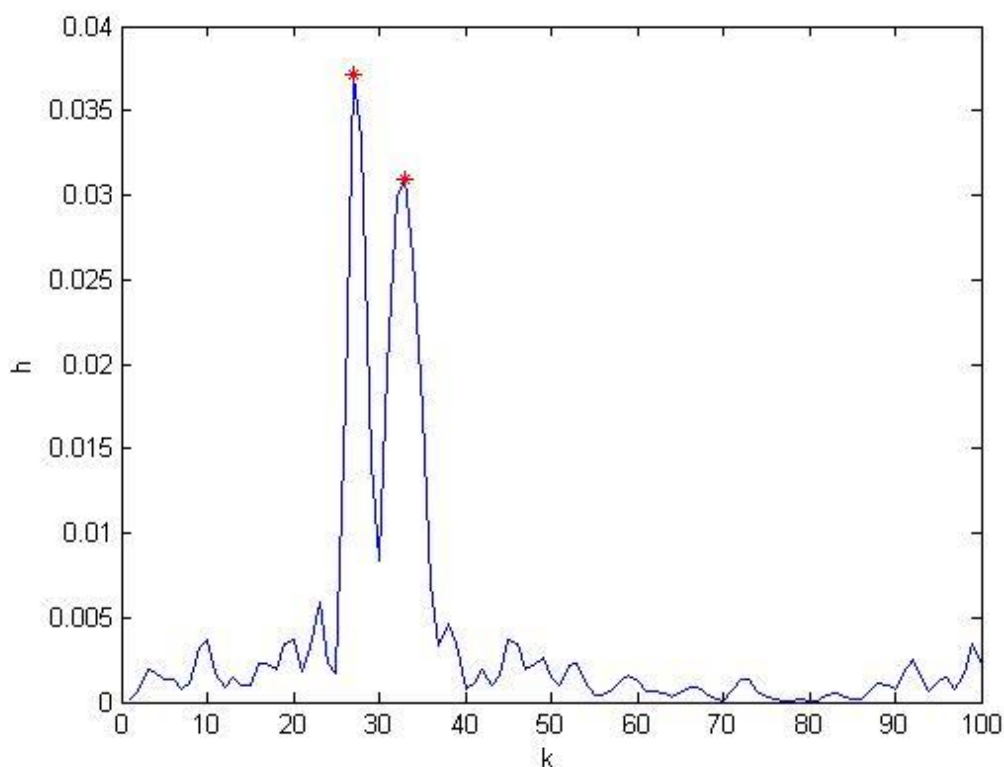


图 3-3

如图 3-3，两个标志了红色星号的两个主峰即为论文定义的主多径。两条主多径的时延估计：

	时延
多径 1	1755 ns
多径 2	2145 ns

3.3.2 多普勒展宽的提取

(1) 公式推导

根据式 (1-8)

$$\begin{aligned}
 r(t) &= \sum_{i=1}^N a_i' e^{-j2\pi \frac{v}{\lambda} t \cos \theta_i} s(t - \tau_i) \\
 &= \sum_{i=1}^N a_i' e^{-j2\pi f_d t} s(t - \tau_i)
 \end{aligned}$$

根据假设 3，发射信号等间隔均匀发射，即发射信号具有周期性，比较不同周期，同一时延的 $s(t)$ ，有：

$$s(t - \tau_i + nT) = s(t - \tau_i + mT), n \neq m$$

进而导出：

$$r(t+nT) = \sum_{i=1}^N a_i s_0 e^{j2\pi f_{di}(t+nT)}$$

其频谱为：

$$R(f) = \sum_{i=1}^N a_i s_0 \delta(f - f_{di})$$

所以，可以对以发射周期为间隔采样的 $r(t)$ 进行 FFT，求得接收信号多普勒谱：

$$FFT(r(t_0) \quad r(t_0+T) \quad \dots \quad r(t_0+NT))$$

(2) 基于快速傅里叶变换（FFT）提取多普勒展宽

类似 3.3.1 中多径数目的提取方法，信道的多普勒展宽都需要对多普勒谱分布进行分析获得。考虑到信道中噪声的影响，为了保证后续统计参数的准确性和提升数据池的精度，数据处理时需要选取一个判决阈值来区分噪声和系统中的有效多普勒频率，进而对数据样本进行筛选。

基本的峰值衰减法确定判决阈值，进行样本筛选时，首先需要确定的两个基本参数是：功率延迟分布的峰值和样本的动态噪底。功率延迟分布的峰值可以通过峰值搜索得到。样本的动态噪底可以通过对信道噪声大小的近似估计得到，计算时，首先确定功率延迟分布中的低幅值点集合（所有采样幅值中最小的 20%）；其次剔除集合中 25% 的最高点和 25% 的最低点（类似中值滤波），并更新集合数据；最后对新集合中的样本点的幅值进行平均，作为动态噪底。峰值衰减算法要求样本功率延迟分布的峰值大于动态噪底 18dB，并将功率延迟分布峰值衰减 15dB 后的门限值作为判决阈值，这样有效样本的判决阈值和动态噪底之间的动态范围应该不小于 3dB。最后，剔除不符合要求的样本。利用剩余的多普勒频谱的点数进行多普勒扩展的估计：

$$Band_{doppler} = N * f_s / M$$

$Band_{doppler}$ 为多普勒展宽， N 为有效多普勒的点数， f_s 为采样频率， M 为 FFT 点数。

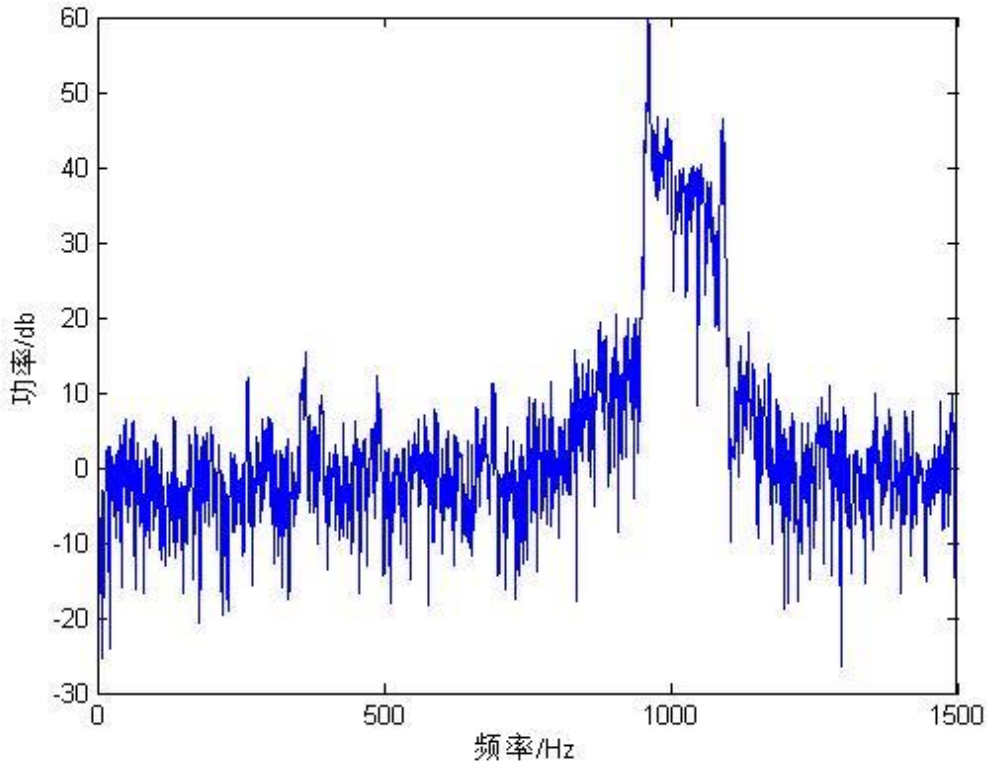


图 3-4 基于 FFT 提取多普勒展宽

如图 3-4，反映了多普勒主峰展宽现象，有效多普勒点数为 159，对应多普勒带宽为：159 Hz。

3.4 无线信道指纹模型评价指标及多径信道指纹模型

基于多径信道的冲激响应以及多普勒展宽等信息，论文建立了多径数目、多径时延、最大多径展宽等特征，主多径平均功率、主多径功率方差以及多普勒展宽等备选信道指纹特征，并且提出了信道指纹模型评价指标：基于三个场景各五个测试样本的“指纹”特征，利用 K-means 聚类进行分类，如果出现两个场景的测试样本特征聚为一类，那么认为特征指标不合理；否则，认为特征指纹模型合理。经过聚类分析，我们得到：当信道指纹选取主多径数、主多径平均功率、多普勒扩展等三个特征时，评价指标显示，信道指纹特征选取较为合理。

3.4.1 K-Means 聚类模型的目标函数

对于给定的一个包含 n 个 d 维数据点的数据集 $X = \{x_1, x_2, \dots, x_i, \dots, x_n\}$ ，其中 $x_i \in R^d$ ，以及要生成的数据子集的数目 K ，K-Means 聚类算法将数据对象组

织为 K 个划分 $C = \{c_k, i = 1, 2, \dots, K\}$ 。每个划分代表一个类 c_k ，每个类 c_k 有一个类别中心 μ_i 。选取欧式距离作为相似性和距离判断准则，计算该类内各点到聚类中心 μ_i 的距离平方和

$$J(c_k) = \sum_{x_i \in c_k} \|x_i - \mu_k\|^2$$

聚类目标是使各类的距离平方和 $J(C) = \sum_{k=1}^K J(c_k)$ 最小。

$$J(C) = \sum_{k=1}^K J(c_k) = \sum_{k=1}^K \sum_{x_i \in c_k} \|x_i - \mu_k\|^2 = \sum_{k=1}^K \sum_{i=1}^n d_{ki} \|x_i - \mu_k\|^2$$

其中， $d_{ki} = \begin{cases} 1, & x_i \in c_i \\ 0, & x_i \notin c_i \end{cases}$ 。

显然，根据最小二乘法和拉格朗日原理，聚类中心 μ_k 应该取为类别 c_k 类各数据点的平均值。

K-means 算法从一个初始的 K 类别划分开始，然后将各数据点指派到各个类别中，以减小总的距离平方和。因为 **K-means** 聚类算法中总的距离平方和随着类别个数 K 的增加而趋向于减小。因此，总的距离平方和只能在某个确定的类别个数 K 下，取得最小值。

3.4.2 K-means 聚类模型的算法流程

K-means 算法是一个反复迭代的过程，目的是使聚类域中所有的样本到聚类中心距离的平方和 $J(C)$ 最小，算法流程包括 4 个步骤，具体流程如图 3-5 所示。

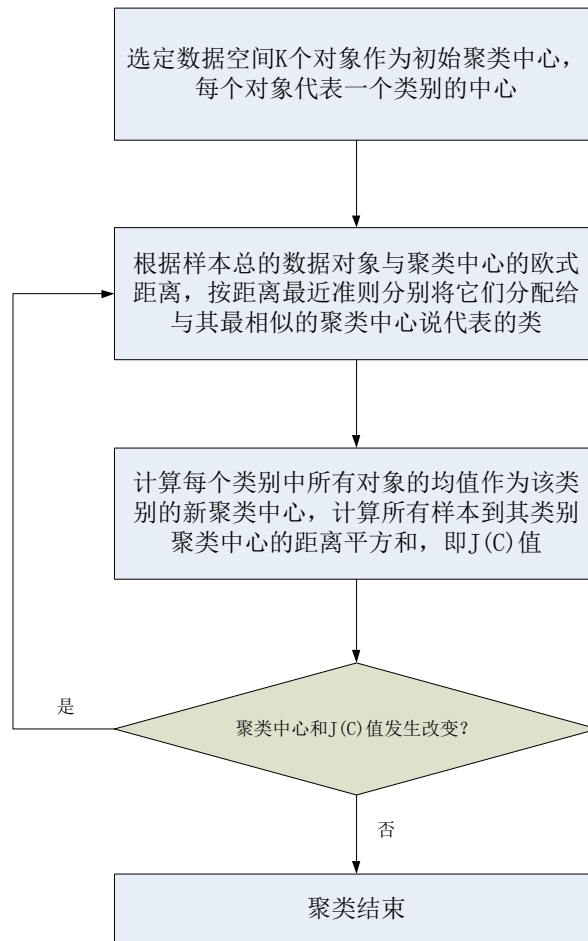


图 3-5

经过 K-means 分析，场景一、二、三测试样本，恰好按照自己的场景聚成一类，聚类结果如图 3-5、3-6 所示：

```

命令行窗口
第1类的有6    7    8    9    10
第2类的有1    2    3    4    5
第3类的有11   12   13   14   15
fx >>
  
```

图 3-6

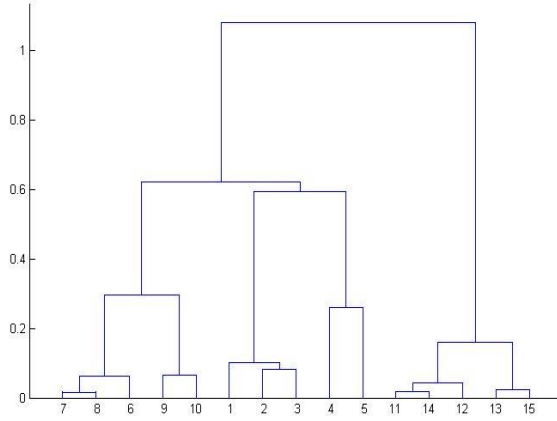


图 3-7

图 3-7 中 1、2、3、4、5 号特征点代表场景一，6、7、8、9、10 号样本点代表场景二，11、12、13、14、15 号特征点代表场景三。从聚类树上可以清晰看到，场景一、场景二、场景三各自的样本点聚为一类，没有发生误判，证明指标选取较好。

主多径：选取

$$h(k) > \frac{1}{2} \max(h(k))$$

$P_{average}$ ：主多径平均功率，即一个场景所有主多径的功率平均。

$$P_{average} = \frac{\sum_{i=1}^N P_{main}(i)}{N}$$

$P_{main}(i)$ ：某个主多径的功率。

N_{main} ：主多径个数。

$Band_{doppler}$ ：多普勒展宽， $Band_{doppler} = N * f_s / M$ 。

针对这一问题，论文提出利用三个场景“指纹”特征利用 K-means 聚类进行分类，如果出现两个场景的测试特征聚为一类，那么认为特征指标不合理；否则，认为特征指纹模型合理。

四、基于多径指纹模型的场景判别

(1) 针对第二个问题，论文基于上一章提出的多径指纹模型，提取了问题二中两个测试场景的指纹特征：

(2) 利用 K-means 聚类，将原有的 15 个特征指纹与新加入的测试指纹进行聚类，通过分类结果，判断第二问中的测试属于哪一个场景。得到结果如下：

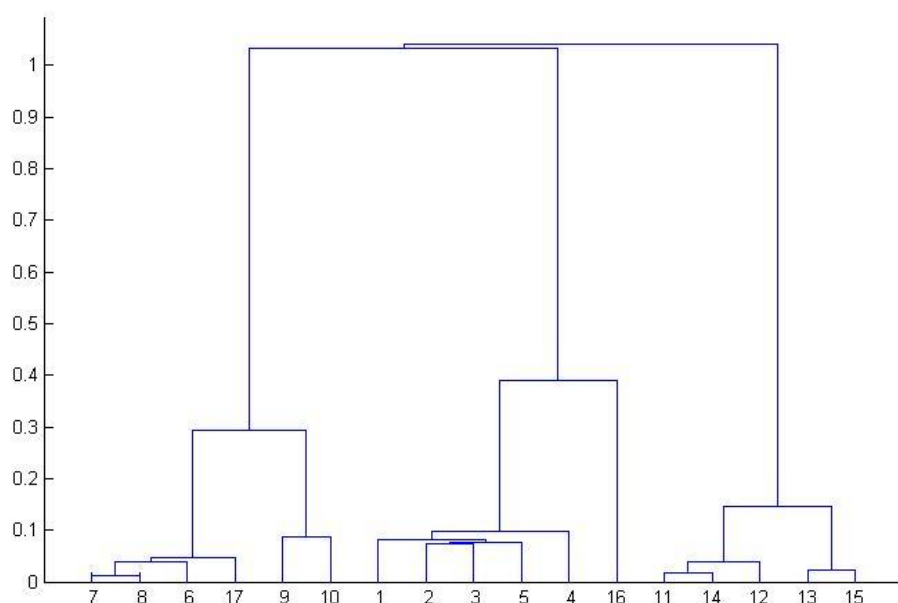


图 4-1

图 中 1、2、3、4、5 号特征点代表场景一，6、7、8、9、10 号样本点代表场景二，11、12、13、14、15 号特征点代表场景三，16 代表测试样本 Test1ForScene.mat，17 代表测试样本 Test2ForScene.mat。从聚类树可以看到，Test1ForScene.mat 测试数据划分在场景一一类中，Test2ForScene.mat 测试数据划分在场景二一类中。

场景一	Test1ForScene.mat
场景二	Test2ForScene.mat

表 4-1

表 4-1 场景与测试样本的对应关系

五、问题三求解

5.1 问题分析

针对问题的要求，即对路段进行分类，因此，本文建立 K-Means 聚类模型用于，先将观测数据进行小间隔、等距离划分，利用问题一建立的模型，计算得到每一小段的多径数目、延迟功率谱、多普勒扩展等“指纹”特征。根据多径数目，先对路段进行粗划分，得到初始分类；然后，将得到的特征参数作为 K-Means 算法的输入，利用“指纹”特征对初始分类进行分层划分，得到最终的

路段分类，并计算每一条路段的“指纹”特征，最后对划分结果是否合理有效进行分析。

5.2 问题求解

5.2.1 求解步骤

步骤一：根据问题一的模型，估计信道模型并利用中值滤波得到该路段的多径数目，并依据多径数目，对路段进行初始划分；

步骤二：在初始划分出来的类别中，以 100 个信道样本为基本单元，一个基本单元的时间长度为 1s，并对每个基本单元进行信道估计，根据估计的信道的系统函数，计算得到每一小段的平均最大延迟幅值以及相应的标准差、时延；

步骤三：将得到的特征参数作为 K-means 聚类模型的输入，在步骤一的基础上，对其进行基于 K-Means 算法的分层聚类，即对划分出来的大类别利用 K-means 聚类算法对基本单元进行聚类，直到每一类内的特征参数无明显差别为止；

步骤四：提取每一个类中的“指纹”特征，对划分的合理性进行分析，判别划分是否合理有效。

5.2.2 路段的初始划分

利用 mopc 算法对信道的系统函数进行计算，得到系统函数，计算每个信道的多径数目，按照多径数目的特点对信道划分，划分结果如图 5-2 所示。

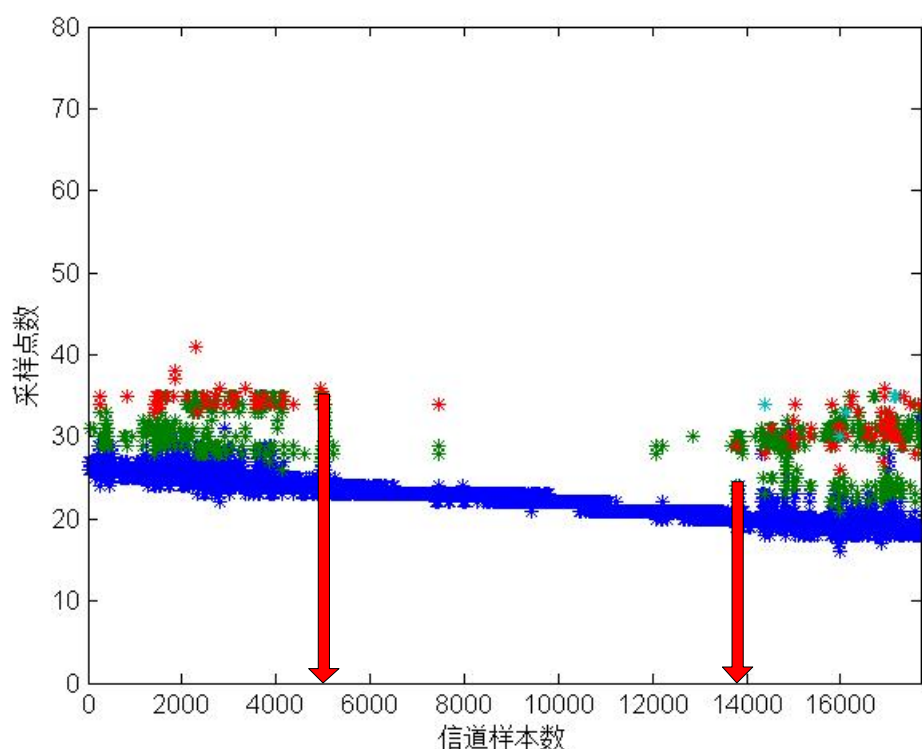


图 5-1

根据图 5-2 可以看出，在 0-5000 范围内每条信道的多径数目大都大于 1，因此将其划分为第一类；在 5000-14000 范围内每条信道的多径数目基本为 1，只存在少数几个大于 1 的，所以将其划分为第二类；在 14000-17700 范围内多径数目大于 1 的信道比较密集，故将其划分为第三类。

5.2.3 基于 K-means 算法的分层聚类

在步骤二的基础上，利用计算得到基本单元的“指纹”特征，按照 K-means 算法的算法流程，对三个大类中的基本单元进行第一次聚类，得到结果如表 5-1 所示。

第一大类	第二大类	第三大类
1-300	5001-10900	14001-15400
301-4200	10901-13000	15401-15700
4201-5000	13001-14000	15701-17700

表 5-1

通过表 5-1 可以看出，第一大类中的第二小类以及第二大类中的第一小类的范围较大，可能包含有多个区域，因此，在这两个区域中基本单元进行聚类，

得到结果如表 5-2 所示。

301-4200	5001-10900
301-2600	5001-5700
2601-4200	5701-9000
	9001-10900

表 5-2

通过观测可以看出，每个类别范围不存在较大差距，最终可得到 12 段，并按照该段划分对其“指纹”特征进行提取，按照如上分类顺序得到结果如表 5-3 所示。

	平均最大延迟功率	平均延迟功率标准差
第一段 1-300	2.3488×10^{-7}	1.2264×10^{-7}
第二段 301-2600	1.8677×10^{-7}	1.2697×10^{-7}
第三段 2601-4200	2.0193×10^{-7}	1.3090×10^{-7}
第四段 4201-5000	2.5821×10^{-7}	1.1757×10^{-7}
第五段 5001-5700	3.4987×10^{-7}	1.3213×10^{-7}
第六段 5701-9000	3.1447×10^{-7}	7.8179×10^{-8}
第七段 9001-10900	2.9153×10^{-7}	8.2970×10^{-8}
第八段 10901-13000	2.7361×10^{-7}	9.0516×10^{-8}
第九段 13001-14000	2.6213×10^{-7}	1.0207×10^{-7}
第十段 14001-15400	1.8247×10^{-7}	1.1257×10^{-7}
第十一段 15401-15700	2.6838×10^{-7}	1.0598×10^{-7}
第十二段 15701-17700	1.4384×10^{-7}	1.0860×10^{-7}

表 5-3

六、问题四的求解

首先，分别计算出两个真实信道的衰落特性，结果如图 6-2、6-3 所示。

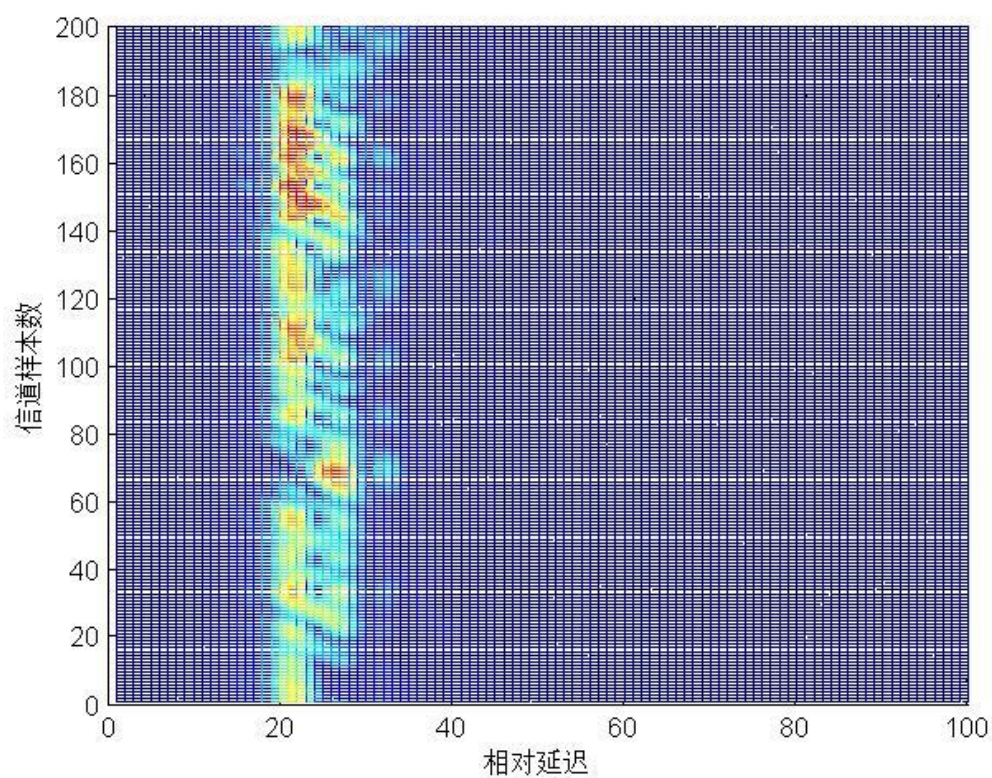


图 6-2

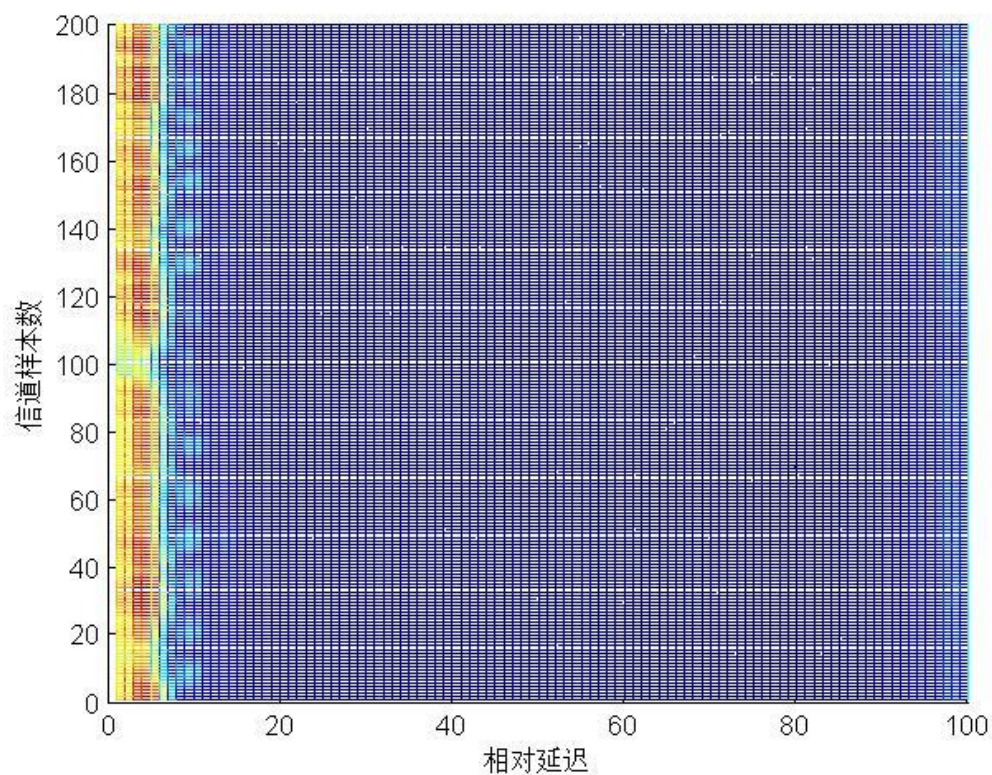


图 6-3

从图 6-2、图 6-3 可以看出，测试信道一明显多径数目大于 1 且平均最大延迟功率不是太大，而在测试信道二中，多径数目主要呈现 1 个且平均最大延迟功率较大。因此，从多径数目的角度，可以得出测试信道二大致处在 5000-14000 这个范围，测试信道一处处在 0-5000 或 14000-17700 这两个范围内。对测试信道一、二分别提取平均最大延迟功率以及平均功率延迟标准差参数，结果如表 6-1 所示。

	测试信道一	测试信道二
平均最大功率延迟	1.2088×10^{-7}	3.3263×10^{-7}
平均功率延迟标准差	6.5681×10^{-8}	6.9692×10^{-8}

表 6-1

将表 6-1 中的数据与表 5-2 中的数据对比可以发现测试信道一比较有可能属于第二段、第十段、第十二段，测试信道二比较有可能属于第六段、第七段。

选取主多径数目、通道平均功率、多普勒扩展带宽作为“指纹”识别特征，运用相关系数法来识别测试信道样本分别属于哪个场景。相关系数是用以反映变量之间相关关系密切程度的统计指标。相关系数是按积差方法计算，同样以两变量与各自平均值的离差为基础，通过两个离差相乘来反映两变量之间相关程度。

根据相关系数的计算公式

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}}$$

可以得出两个测试信道与其可能相似的路段的相关矩阵，相关矩阵各位置的数值对应相应行列的相关系数。

最后，根据计算出来的相关系统可以发现测试信道一处处在路段的 2.5m 到 21.7m，测试信道二处在路段的 47.5m 到 75m。

七、模型评价及改进

7.1 模型优点

1、论文基于 MPOC 算法，求解了多径信道的单位脉冲响应，并且基于多径信道的单位脉冲响应，提取了多径数目、多径时延、最大多径展宽等特征，相比于传统扩频滑动相关法、最小二乘法具有更高的时延分辨力。

2、在多径检测与多普勒展宽有效点提取的时候，本文用数据中较小的点求平均构造噪声基底作为检测门限，取过噪声门限作为有效数据，可以防止噪声数据被作为有效数据使用，起到自适应门限设置的效果。

3、论文针对求解出的多径信道数量较多且除少数主峰以外，其余多径的增益系数很小，很难区分是多径还是主峰引起的旁瓣或噪声，；论文提出了取最大主峰值幅值 0.5 倍以上的峰值作为主多径来稳定提取多径特征的方法，属创新。

7.2 模型缺点

1、论文提出采用主多径这一概念来使多径特征稳定，但是论文中取最大主峰值幅值 0.5 倍以上的峰值作为主多径的方法缺乏理论支撑，属经验做法。

2、由于时间关系，论文未从接收信号是否具有直达径这个角度进行分析。

八、参考文献

- [1]张峻铭，无线信道多径时延估计及信道建模：[硕士学位论文]. 成都. 电子科技大学. 2013. 4-6
- [2]张峻铭，无线信道多径时延估计及信道建模：[硕士学位论文]. 成都. 电子科技大学. 2013. 14-16
- [3]高林毅，室内宽带无线信道测量与建模技术研究：[硕士学位论文]. 北京 . 北京邮电大学. 2011 . 12-13
- [4]赵敏，高速铁路无线宽带信道测量、数据分析与建模研究：[博士学位论文] 北京邮电大学. 67-68
- [5]司守奎 孙玺菁，数学建模算法与应用，北京，国防工业出版社. 2014